Towards precision genomics

Opportunities for CSER

Euan Ashley MRCP, DPhil

Associate Professor of Cardiovascular Medicine & Genetics

Director, Center for Inherited Cardiovascular Disease

Co-Director, Clinical Genomics

Chair, Biomedical Data Science Initiative

Stanford MEDICINE

Founder and advisor

Advisor, Academic grant

Advisor,
Clinical trial site PI

Academic grant,
Clinical trial site PI

Academic grant

In kind grant support

In kind grant support

In kind grant support

Academic grant

The First Child Saved By DNA Sequencing

+ Comment Now + Follow Comments

THE NEW YORKER

MEDICAL DISPATCH | JULY 21, 2014 ISSUE

ONE OF A KIND

What do you do if your child has a condition that is new to science?

BY SETH MNOOKIN

Genomic medicine is here

. . . for rare disease

and NIPT

Genome study solves twins' mystery condition

Sequencing ends years of speculation over children's rare disorder.

Erika Check Hayden

Two years ago, 13-year-old Alexis Beery developed a cough and a breathing problem so severe that her parents placed a baby monitor in her room just to make sure she would survive the night. Alexis would often cough so hard and so long that she would throw up, and had to take daily injections of adrenaline just to keep breathing. Yet doctors weren't sure what was wrong.

Genome sequencing suggested a new approach to treatment for twins Noah and Alexis Beery, shown here with their parents.

Life Technologies

HealthlineNews

Healthline → Healthline News → Undiagnosed Diseases Program Ends Mother's 20-Year Search for Answers

Undiagnosed Diseases Program Ends Mother's 20-Year Search for Answers

Written by Sandra Levy | Published on July 10, 2014

When the lights finally came back on after Hurric
Samantha Anastasia was ecstatic, but it wasn't j
neighbors were finally out of the dark. After 20 y
two of her three children, she had just returned f
(NIH) Undiagnosed Diseases Program (UDP) in
finally learned what had caused two of her childr
walk when they were just infants.

So when will we get personalized medicine for everyone?

Genome Seen As Medical Crystal Ball

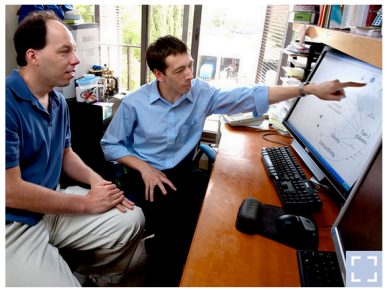
APRIL 30, 2010 12:01 AM ET

RICHARD KNOX

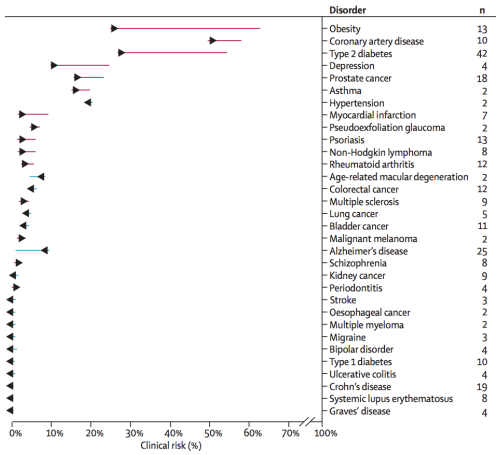
npr

Listen to the Story
Morning Edition

Stanford scientist Steve Quake was only the fifth person in the world to have his entire genetic code -- his genome — spelled out last summer. Now he claims to be the first to use it to find out just what diseases he's at risk for, and what he should do about it.

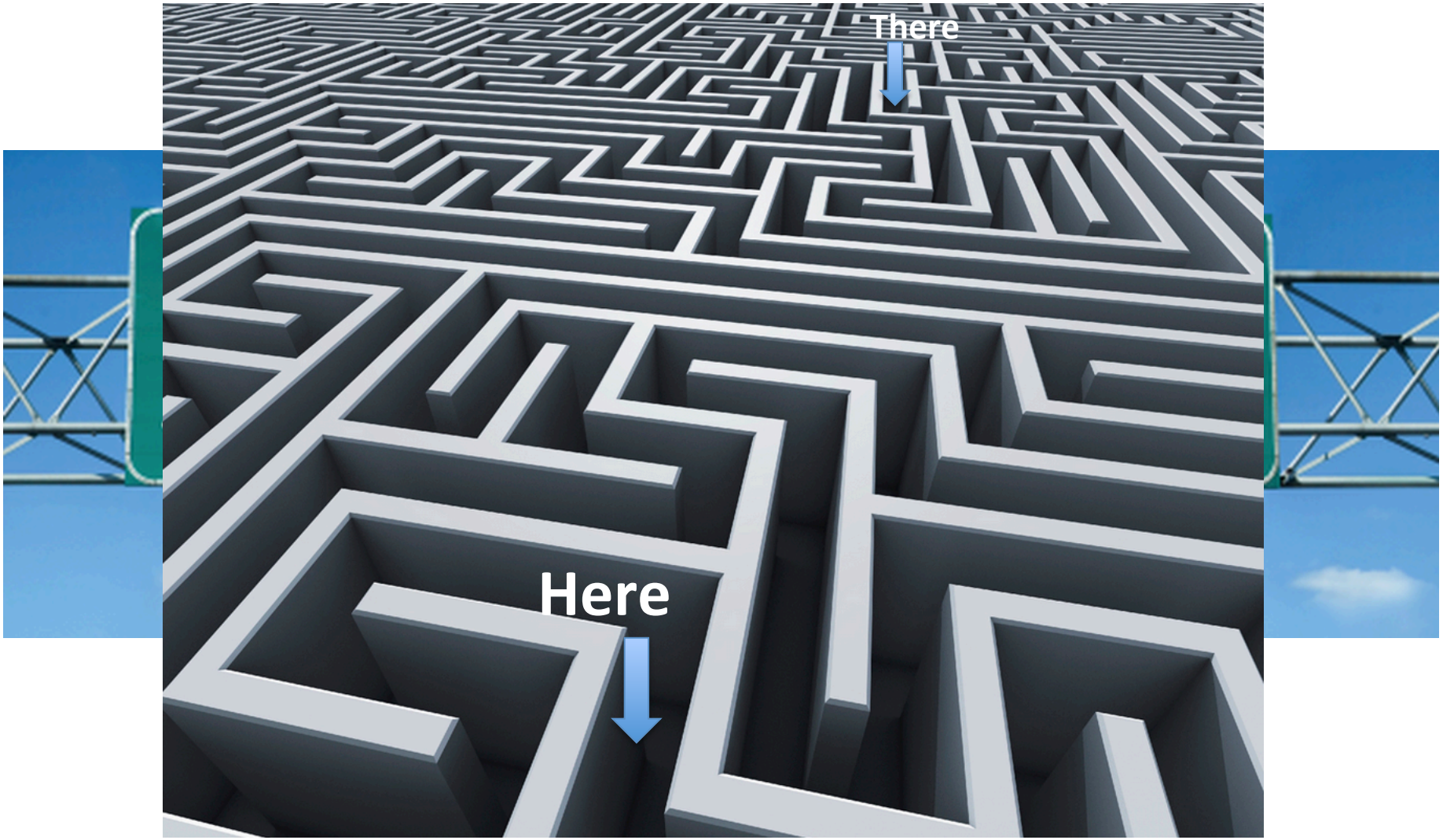


Stephen Quake (left) and Euan Ashley review genetic risk factors that Ashley and colleagues found in Quake's genome.



Lancet 2010; 375: 1525–35

Most people overestimate what they can do in one year and underestimate what they can do in ten years.


There
Here

Expanding clinical utility

- Choose use cases
- Design, implement, test, repeat
- Build evidence base for effectiveness
- Assess cost effectiveness
- Include payors early

CHALLENGES

Tommy Caldwell ascending the Dawn Wall of El Capitan
Brett Lowell: National Geographic

1. The genome is complex

2. You can’t call it, if you can’t see it

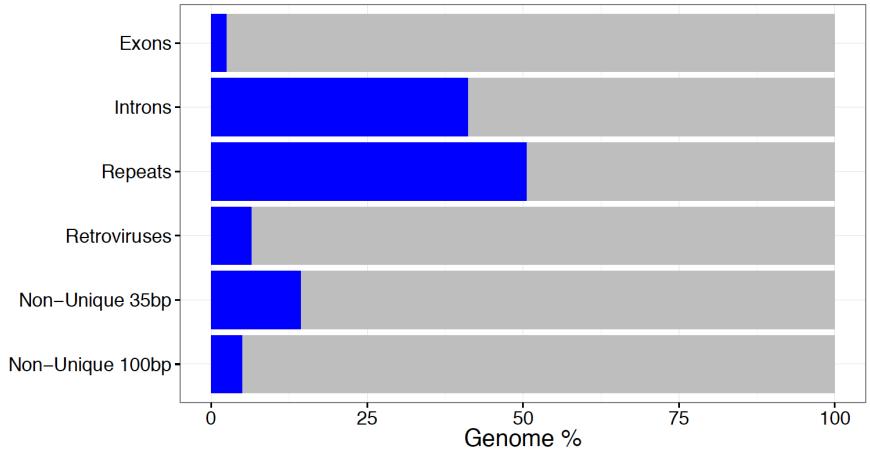
3. The technical performance of our calling algorithms was optimized for cohort variant discovery, not n=1

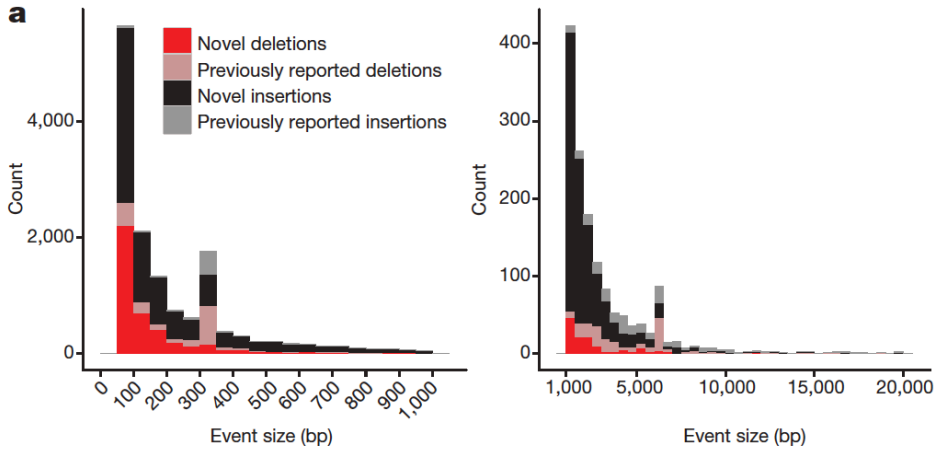
Repeats = ~56% of the genome

Paralogous sequence

- Segmental duplications
- Gene families
- Pseudogenes
  - ~8k
  - Varying constraint

No consensus over how to handle multiply mapped short reads



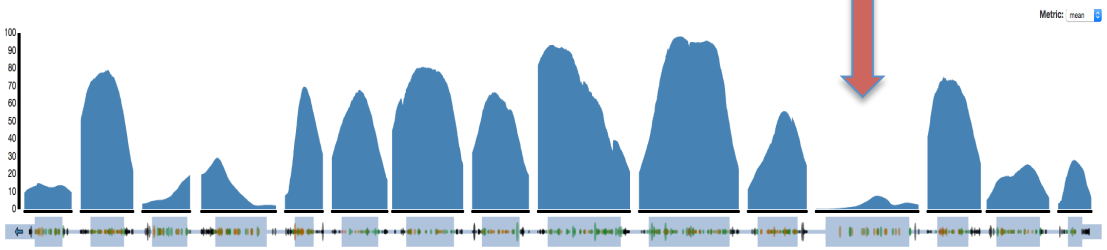


Genome Medicine (in revision)
Nature 517:608 2015

Exome

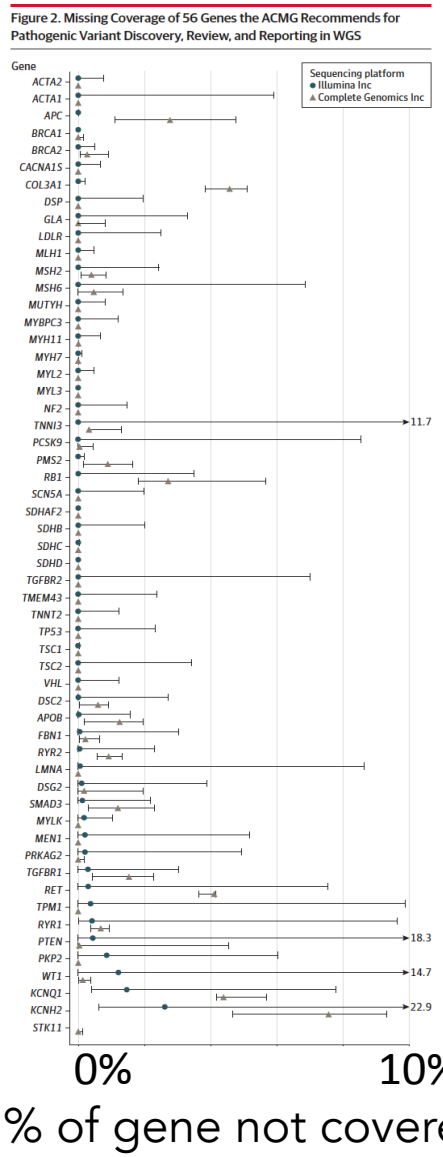
KCNH2

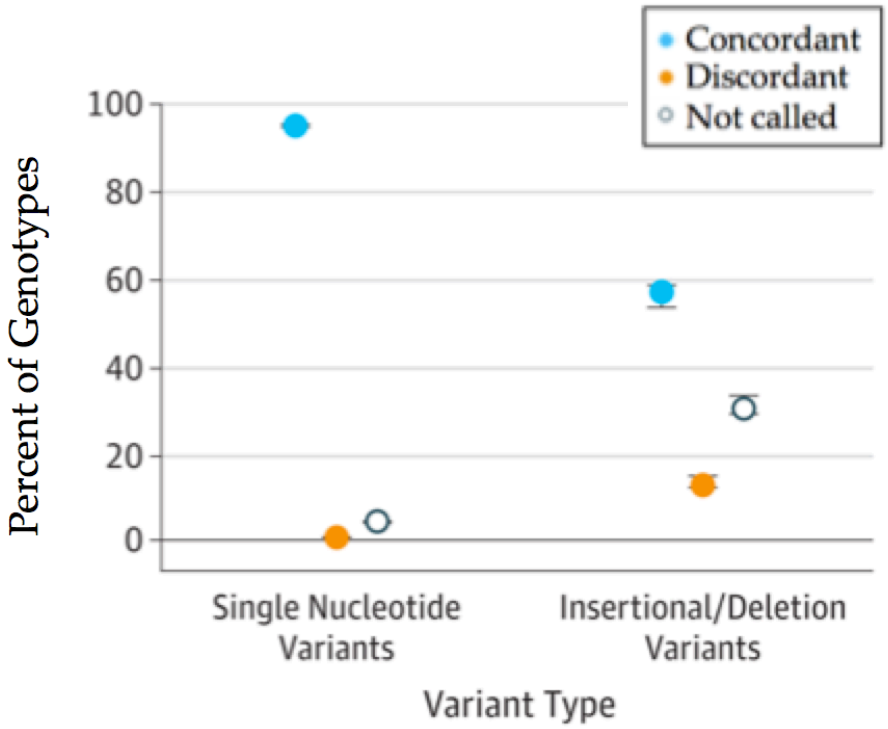
19 pathogenic or likely pathogenic variants in Clinvar

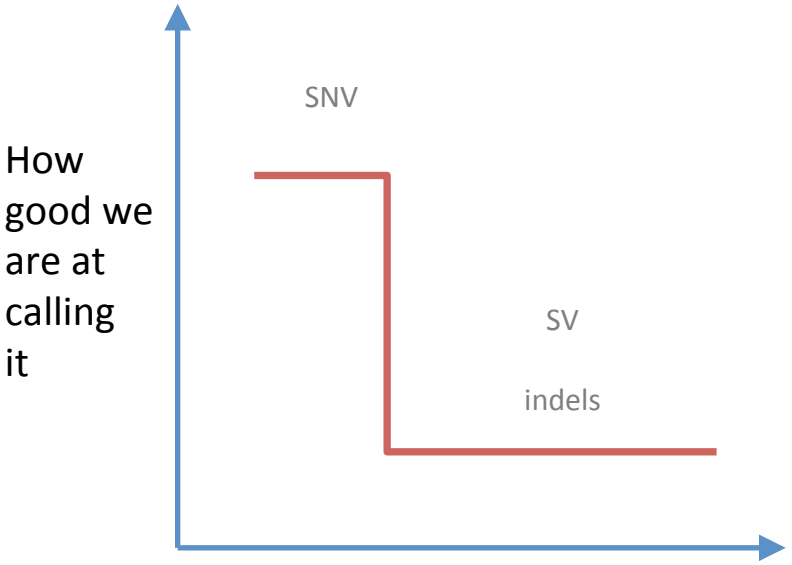


exac.broadinstitute.org

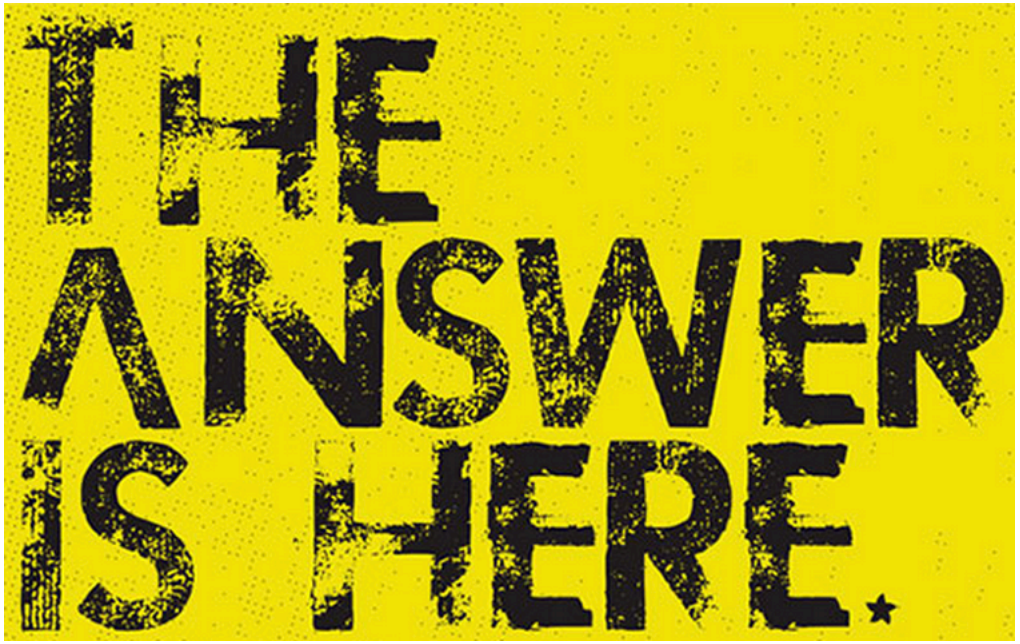
Genome







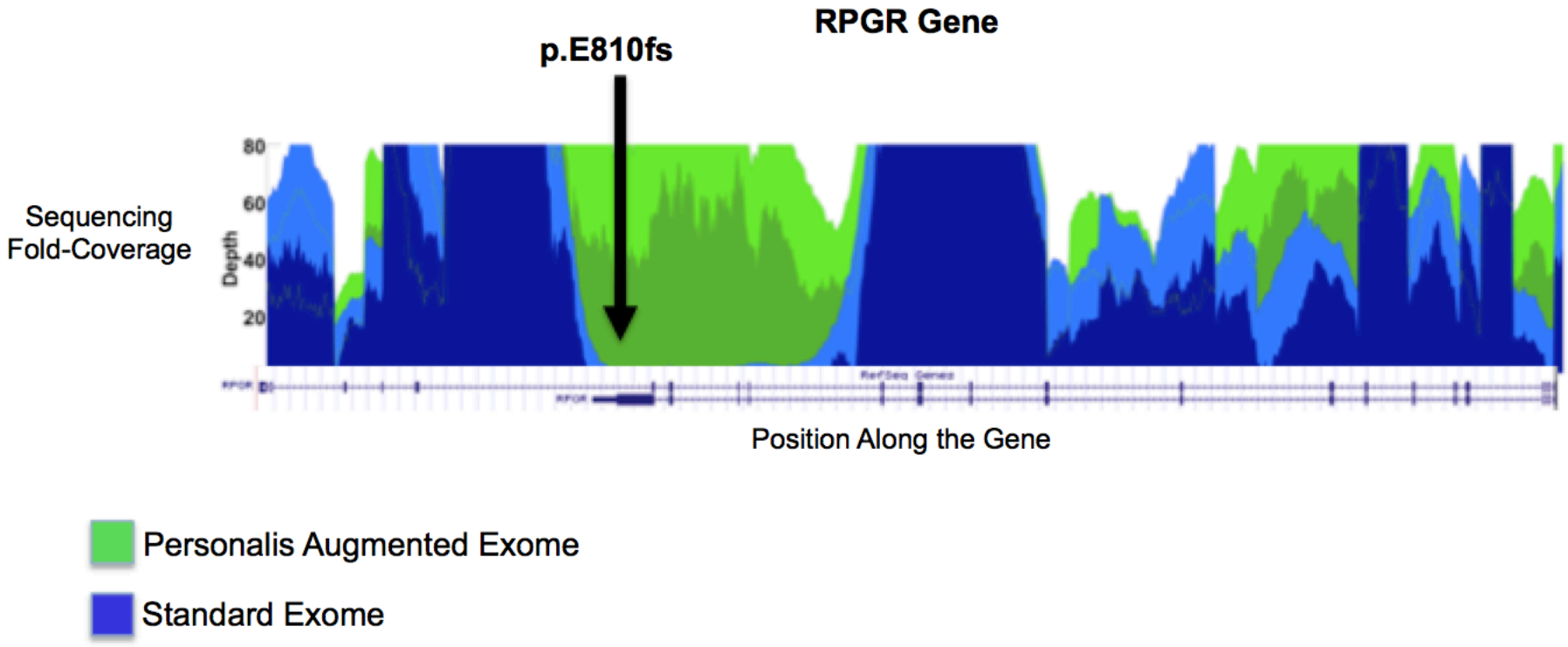
JAMA 2014 311:1035-44 2014



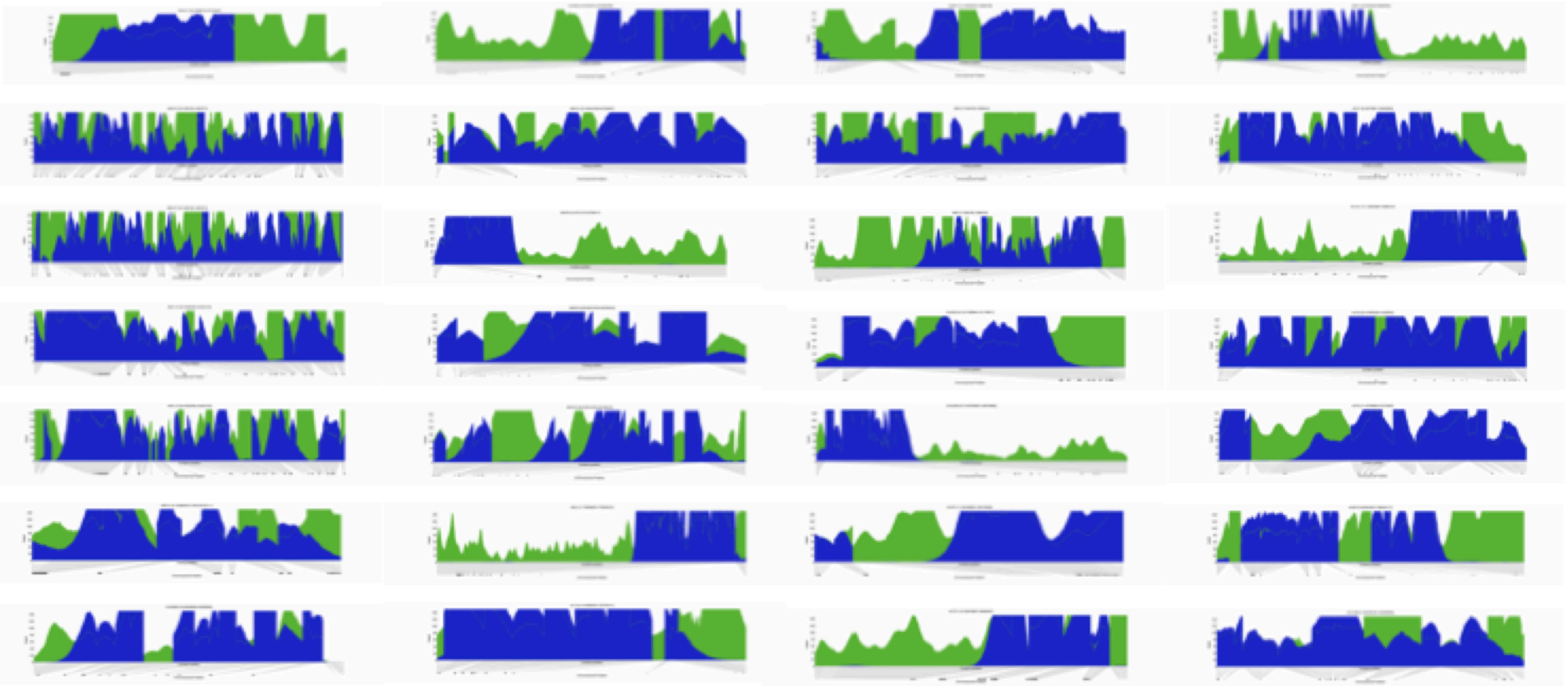
SO WHAT IS THE ANSWER

And does it start with an “E” or a “G”?

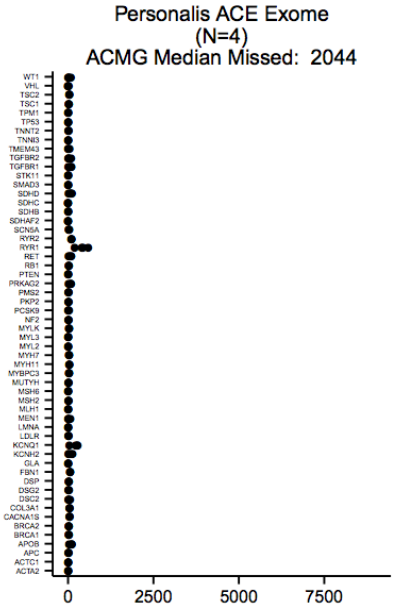
GET COVERAGE

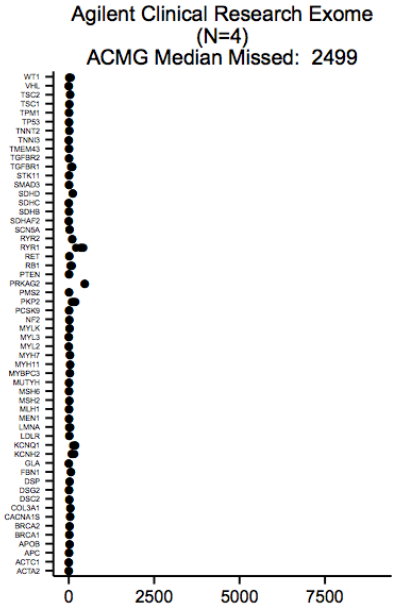


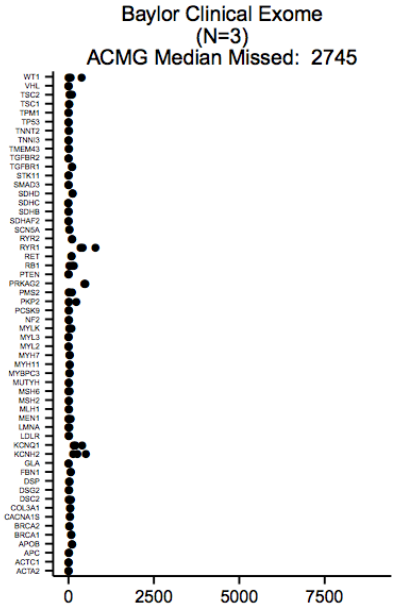
Genome Medicine **7**:71 (2015)

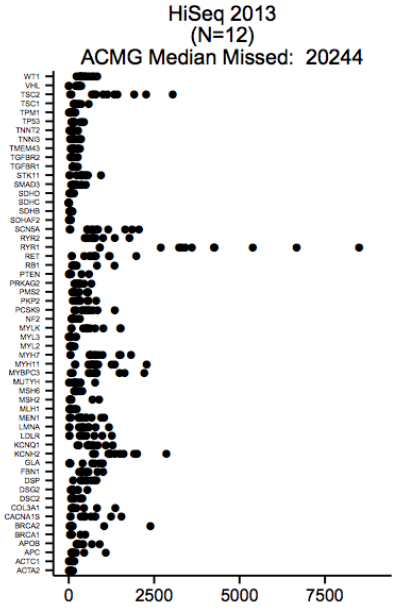


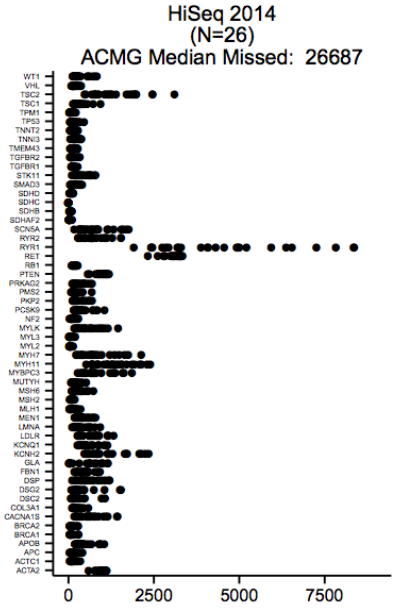
Number of bases not covered by >=20 Q30 bases







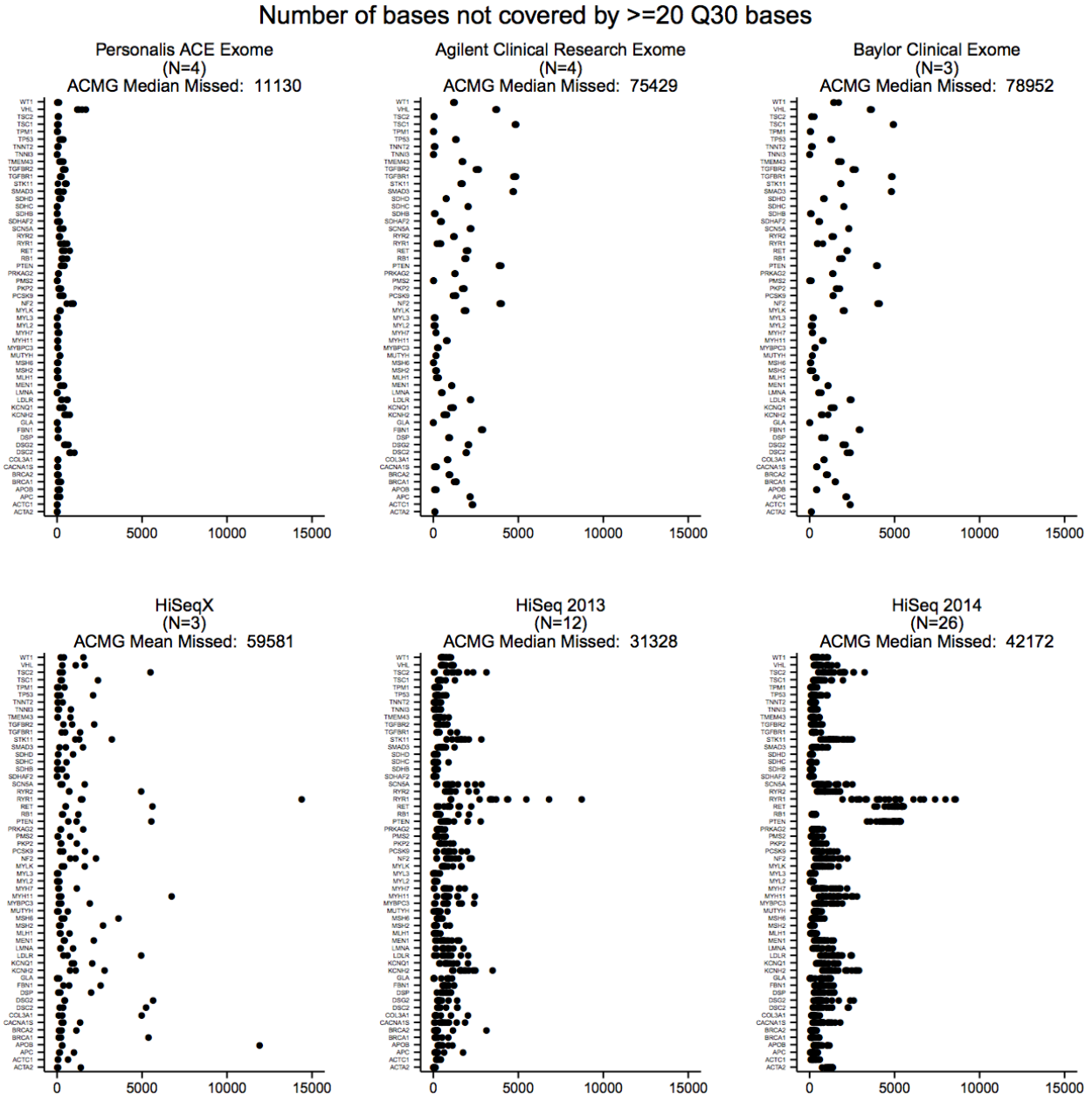




	Gbp
Illumina HiSeq X	111.11
Illumina Hiseq 2013	151.28
Personalis ACE	12.74
Agilent Clinical Research Exome	13.49
Baylor	11.32

Rachel Goldfeder

Exons (includes UTRs)



Rachel Goldfeder

MOCK REPORT - Stanford Clinical Genomics Service

300 Pasteur Drive, Stanford, CA, 94305

Patient Name:	"CGS-3"	Specimen type:	Peripheral Blood
DOB:	35 years old	Date specimen obtained:	3/13/2014
Lab Accession:	PG0001569-BLD	Date specimen received:	3/14/2014
Sex:	Female	Referring physician:	Dr. Euan Ashley
Race:	Caucasian	Genetic Counselor:	Colleen Caleshu
		Referring facility:	Stanford Center for Inherited Cardiovascular Disease

TEST PERFORMED – Genome Sequencing

INDICATION FOR TEST - Clinical diagnosis of hypertrophic cardiomyopathy (HCM) and family history suspicious for sudden cardiac death.

RESULT: Negative – Established or likely causes of the reported phenotype were not identified

INTERPRETATION SUMMARY:

Genome sequencing and variant analysis did not identify an established or plausible explanation for hypertrophic cardiomyopathy (HCM) in this individual. Genes with an established or likely role in HCM and related pathways were analyzed, and 99.91% of the base pairs in the coding regions of these genes were covered by at least 10 independent reads. We did not identify any likely pathogenic variants in genes with a plausible role in HCM. Up to 3.67% of the coding region of the following cardiomyopathy-related genes were not fully covered by at least 10 independent reads, and therefore may be more prone to false negative results: *LDB3, LMNA, HRAS, KCNH2, KRAS, ABCC9, CRYAB, ILK, KCNQ1, PDLIM3, MYH6, DSP, RYR2, DES, PKP2, GATAD1, PRKAG2, TTN, LAMP2, NEBL, DTNA, EMD, EYA4, JPH2, JUP, MYBPC3, RBM20, SOS1, TAZ*. The total number and percent of coding base pairs not covered by at least 10 base pairs are summarized in the table at the end of this report.

Please note that mitochondrially encoded genes are not examined by this test at this time. Consequently the following mitochondrially encoded genes associated with various forms of cardiomyopathy in the scientific literature are not analyzed with this test due to limitations in variant calling methodologies at this time: *MT-TK, MT-TL1, MT-TG, MT-TS1, MT-TH, MT-ND1, MT-TQ, MT-TI, MT-TS2, MT-TD, MT-TL2, MT-ND5, MT-TM, MT-ND6*. This patient has previously had a 18 gene HCM genetic testing panel at GeneDx in March 2012, which included full sequencing of 4 of these 10 mitochondrial genes: *MT-TG, MT-TI, MT-TK, MT-TQ*. If a mitochondrial gene abnormality is highly suspected, consideration of additional testing or reanalysis by this service in the future should be considered to capture the 10 remaining mitochondrial genes.

RECOMMENDATIONS:

1. Clinical correlation is recommended.

2. It is recommended that any 1st degree relative receive continued clinical evaluation and follow-up for features of HCM.

3. Genetic counseling is recommended for this individual and the family.

4. A medical provider can request reanalysis of the genome data, and this is recommended on an annual basis. Data from this genome sequencing analysis can be reassessed for the presence of any variants that may be newly linked to established genes or to newly characterized genes and/or disorders identified since the date of this report that could be associated with the patient's phenotype, based on currently available scientific information. Please contact the laboratory for more information and charges at the time reanalysis is requested.

TEST METHOD: Genomic DNA was extracted from the submitted specimen, a library was generated using the Illumina TruSeq DNA PCR-Free Sample Preparation Kit, and genome sequencing was performed using the HiSeq 2500 System with 100 bp paired-end reads. The DNA sequence was mapped to, and analyzed in comparison with, the published human genome build hg19, Feb 2009 (GRCh37), using Stanford MedGAP v2.0. The reference genome and exome (RefSeq gene model, NCBI Reference Sequence Database, Exons only, April 2014) were assessed for the average depth of coverage and data quality threshold values*.

*The values below represent metrics from this individual's genome sequencing. All values are calculated at base phred quality ≥20.

Mean Depth of Coverage	32.10X
% of genome covered at ≥10X	97.3%
% of exome covered at ≥10X	95.1%

LIMITATIONS:

Absence of a plausible explanation for the reported phenotype by genome sequencing does not exclude a genetic basis of the patient's condition. Some types of genetic abnormalities (e.g., insertions or deletions >10bp, copy number changes, structural variants, and trinucleotide repeat expansions) may not be detectable with the technologies performed by this genome analysis test. It is possible that the genomic region where a pathogenic variant exists in the proband was not covered using the current technologies and therefore was not detected. Additionally, it is possible that a particular genetic abnormality may not be recognized as the underlying cause of the genetic disorder due to incomplete scientific knowledge about the function of all genes in the human genome and the impact of variants in those genes. Only variants in genes associated with the medical condition, or thought to be potentially clinically relevant for the proband's medical condition, are reported here.

The below table, derived from this individual's genome sequencing data, lists coverage metrics for genes known to be associated with HCM. Bases covered by at least 10 reads with a base quality phred score ≥20 at that position are associated with high confidence in detection of heterozygous variants. Bases covered by less than 10 reads with a base quality phred score ≥20 at that position are at increased risk for false-negative variant calls.

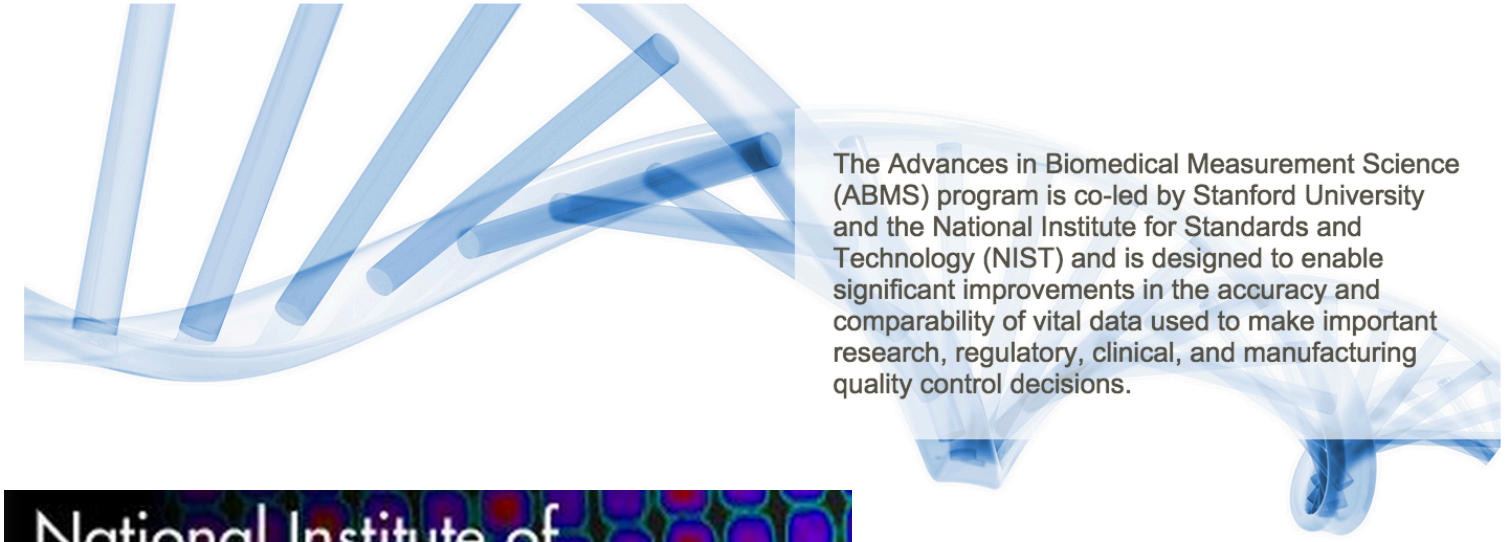
Gene	Total number of coding base pairs	Total number of coding base pairs not covered by at least 10 reads	Percentage of coding base pairs not covered by at least 10 reads
LDB3	6504	109	1.676
LMNA	4085	47	1.151
HRAS	1227	45	3.667
KCNH2	5448	32	0.587
KRAS	5656	26	0.46
ABCC9	8423	24	0.285
CRYAB	1079	22	2.039
ILK	2173	17	0.782
KCNQ1	3370	15	0.445
PDLIM3	3031	13	0.429
MYH6	5902	10	0.169
DSP	9706	7	0.072
RYR2	16260	7	0.043
DES	2239	6	0.268
PKP2	4425	5	0.113
GATAD1	4592	4	0.087
PRKAG2	3966	3	0.076
TTN	113875	3	0.003
LAMP2	9378	2	0.021
NEBL	9894	2	0.02
DTNA	10699	1	0.009
EMD	1333	1	0.075
EYA4	5684	1	0.018
JPH2	5934	1	0.017
JUP	3494	1	0.029

KNOW THE ENEMY

Advances in Biological and Medical Measurement Science

Log in with WebAuth

Home | Program Components | People | ERCC 2.0 Workshop | Genome in a Bottle Consortium | Membership



The Advances in Biomedical Measurement Science (ABMS) program is co-led by Stanford University and the National Institute for Standards and Technology (NIST) and is designed to enable significant improvements in the accuracy and comparability of vital data used to make important research, regulatory, clinical, and manufacturing quality control decisions.

Stanford University

GIAB Reference Materials and Data

The Genome in a Bottle Consortium has selected several genomes to produce and characterize as reference materials. The National Institute of Standards and Technology (NIST) is developing NIST Reference Materials from these genomes, which are DNA extracted from a large homogenized growth of B lymphoblastoid cell lines from the Coriell Institute for Medical Research. Note that there may be small differences between the NIST DNA and the Coriell DNA since they come from different growths of cells, thoughwe do not expect these differencess to be large for most applications.

Read more

Recent Blog Posts

- Preprint describing GIAB PGP data now on biorxiv
- The pilot GIAB/NIST Reference Material 8398 is now available!
- Presenting about GIAB at conferences, abstract, and slide deck

FDA to develop precision-medicine platform

By Steven Ross Johnson | August 6, 2015

The Food and Drug Administration is developing an open-source, cloud-based software platform that will allow for collaborative sharing of information among genomic researchers as part of President Barack Obama's Precision Medicine Initiative.

The agency plans to release a beta version of the platform, named precisionFDA, by December. The goal is that the software allows users to create an informatics community where researchers can store and share their work with collaborators.

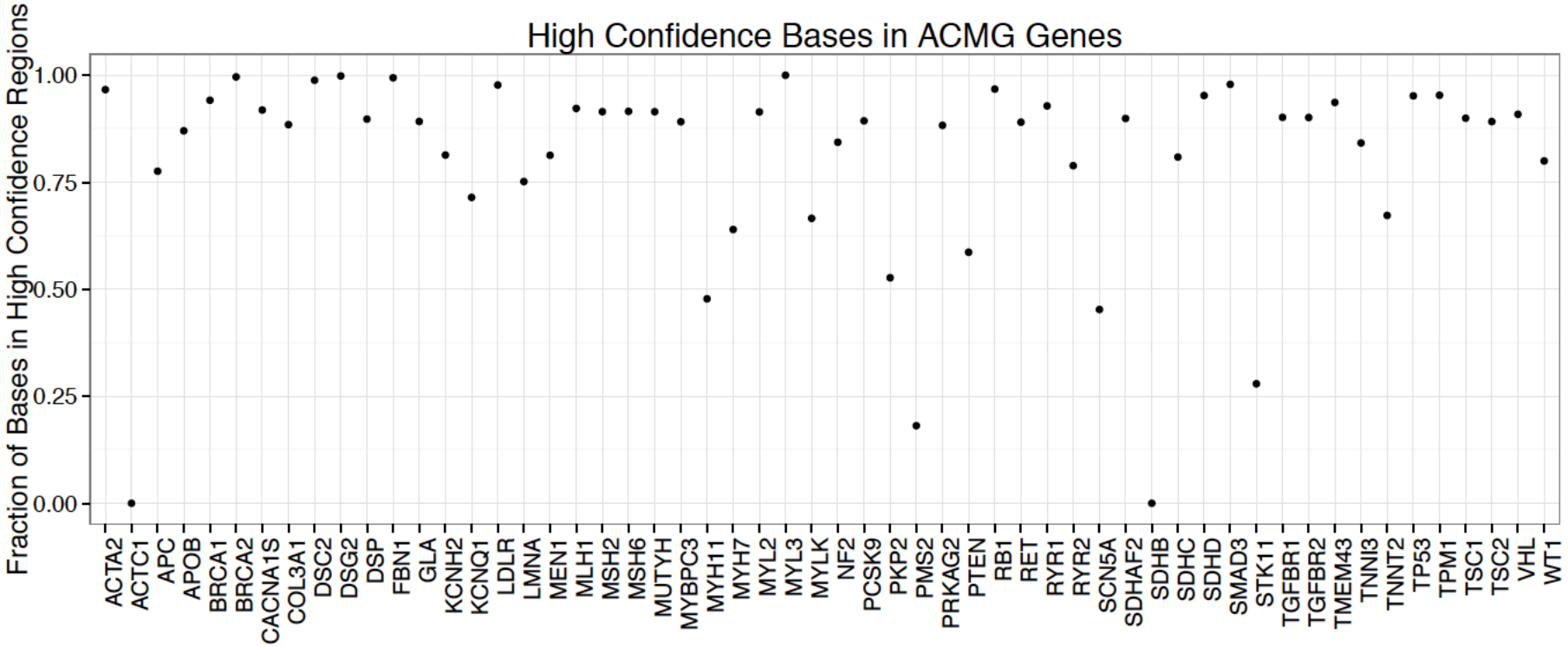
"To begin to realize this new vision, precisionFDA is designed as a crowd-sourced, cloud-based platform to advance the science needed to develop the necessary standards. PrecisionFDA will supply an environment where the community can test, pilot, and validate new approaches."

http://goo.gl/UOcQc8

FDA

DNAnexus

Are the things that matter most in high confidence regions?



NA12878
Clinvar
pathogenic
SNVs

bioRxiv beta
THE PREPRINT SERVER FOR BIOLOGY

Extensive sequencing of seven human genomes to characterize benchmark reference materials

Justin M Zook, David Catoe, Jennifer McDaniel, Lindsay Vang, Noah Spies, Arend Sidow, Ziming Weng, Yuling Liu, Chris Mason, Noah Alexander, Dhruva Chandramohan, Elizabeth Henaff, Feng Chen, Erich Jaeger, Ali Moshrefi, Khoa Pham, William Stedman, Tiffany Liang, Michael Saghbini, Zeljko Dzakula, Alex Hastie, Han Cao, Gintaras Deikus, Eric Schadt, Robert Sebra, Ali Bashir, Rebecca M Truty, Christopher C Chang, Natali Gulbahce, Keyan Zhao, Srinka Ghosh, Fiona Hyland, Yutao Fu, Mark Chaisson, Jonathan Trow, Chunlin Xiao, Stephen T Sherry, Alexander W Zaranek, Madeleine Ball, Jason Bobe, Preston Estep, George M Church, Patrick Marks, Sofia Kyriazopoulou-Panagiotopoulou, Grace Zheng, Michael Schnall-Levin, Heather S Ordonez, Patrice A Mudivarti, Kristina Giorda, Marc Salit, Genome in a Bottle Consortium

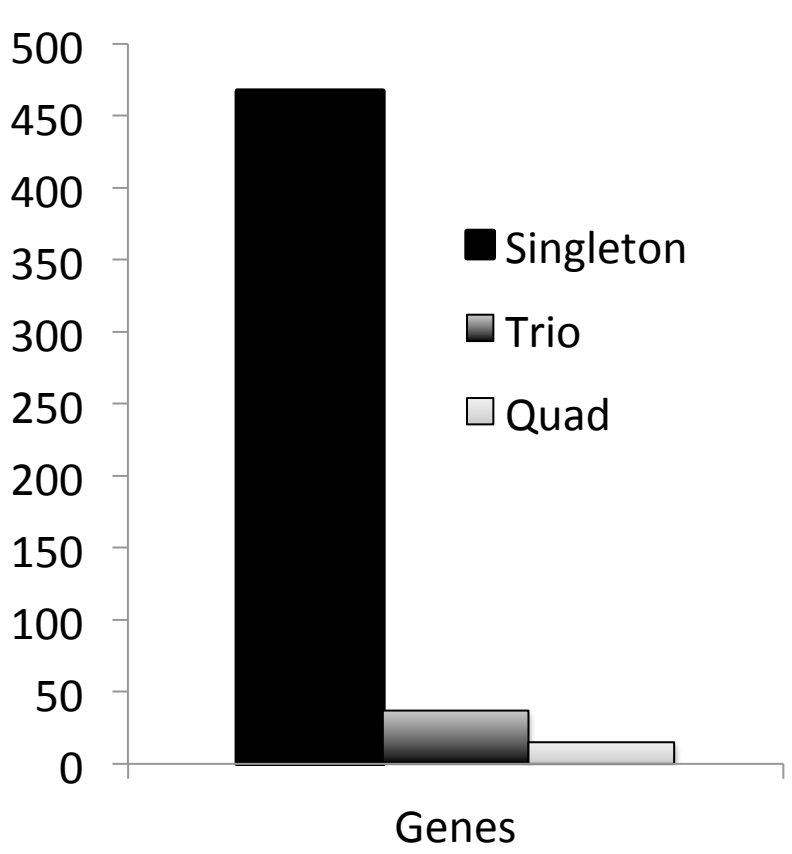
doi: http://dx.doi.org/10.1101/026468

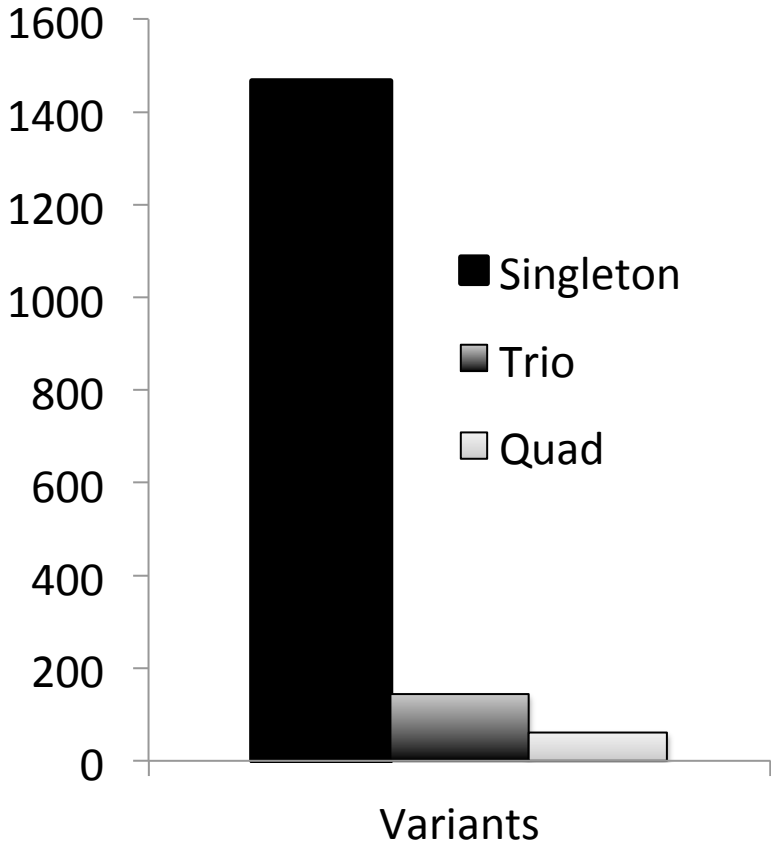
Sample	Son	Father	Mother
Total PASS	55	78	56
Autosome	36	60	34
Recessive	17	10	13
Compound Het	2 genes (2818bp, 340bp diff)	1 gene (340bp diff)	2 genes (2818bp, 340bp diff)
Structural Variants	1DEL (1 variant) Diff End Breakpoint Than Mother	2DEL (2; 20 variants)	1DEL(2 variants)
Mitochondrial	19	18	22
Low Frequency MT Heteroplasmy	10	9	10
ACMG Genes	3	3	1

Posted September 15, 2015.

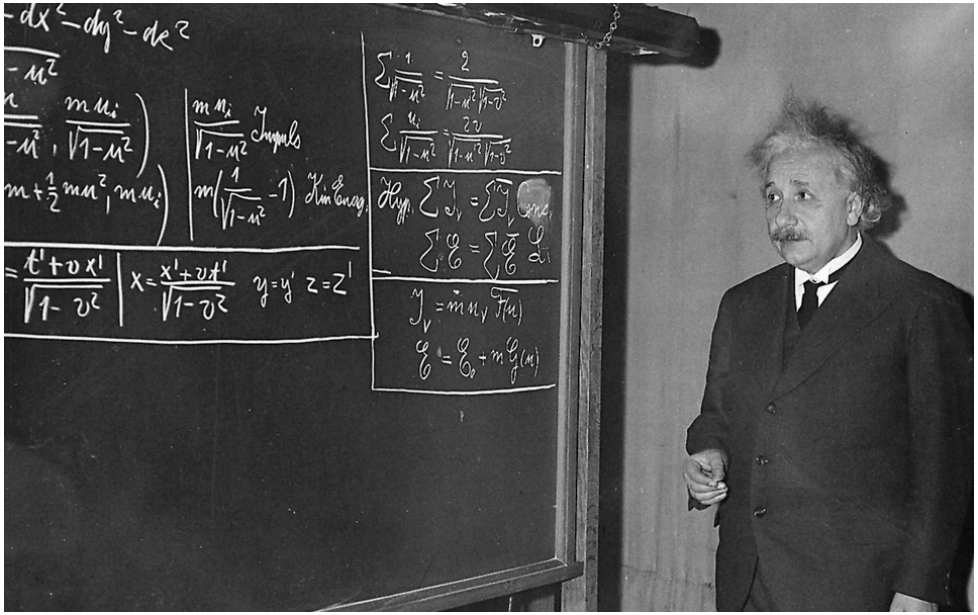
KEEP IT IN THE FAMILY

The strength of family analysis

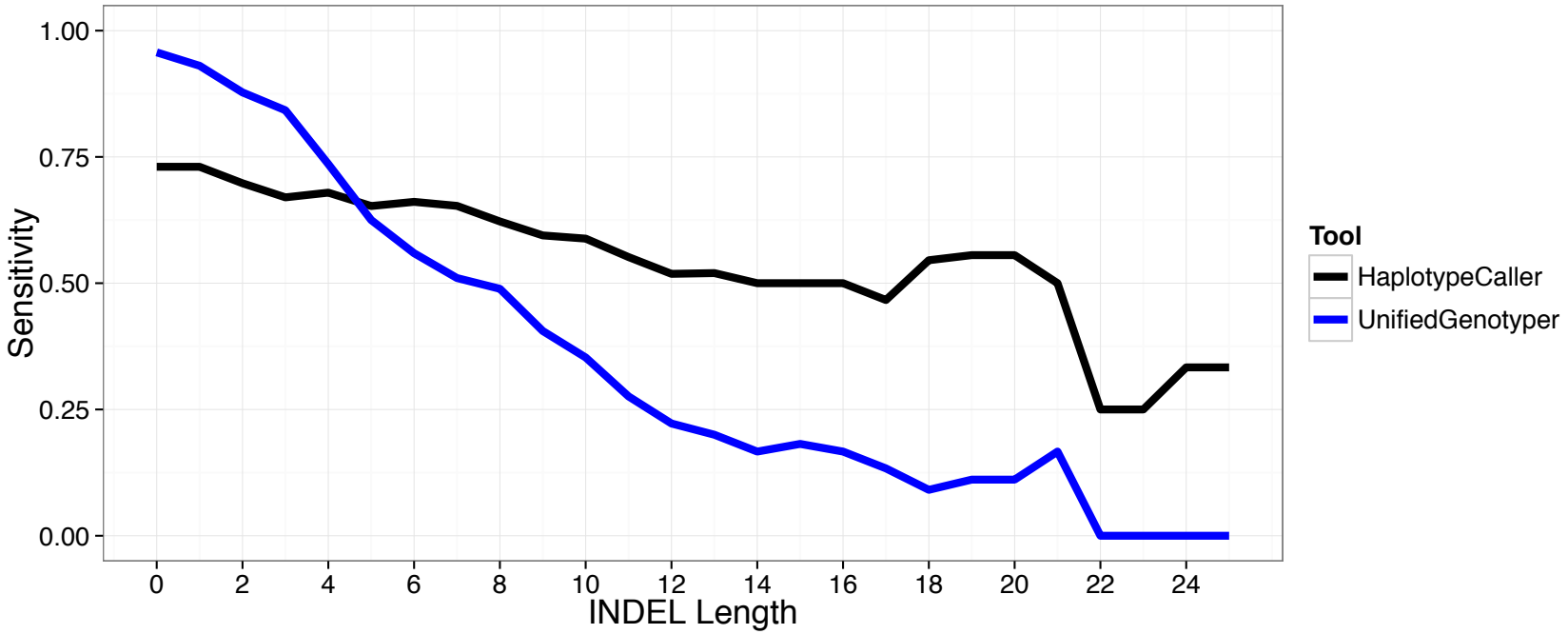




Megan Grove, Dianna Fisk
Compound heterozygosity model



DESIGN BETTER ALGORITHMS



Rachel Goldfeder
*chr22, simulated data

Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications

Andy Rimmer[1,5], Hang Phan[1,5], Iain Mathieson[1], Zamin Iqbal[1], Stephen R F Twigg[2], WGS500 Consortium[3], Andrew O M Wilkie[2], Gil McVean[1,4] & Gerton Lunter[1]

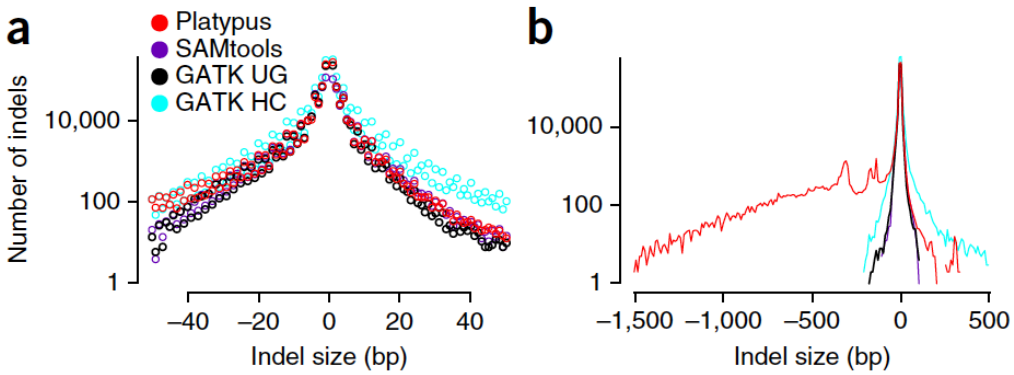
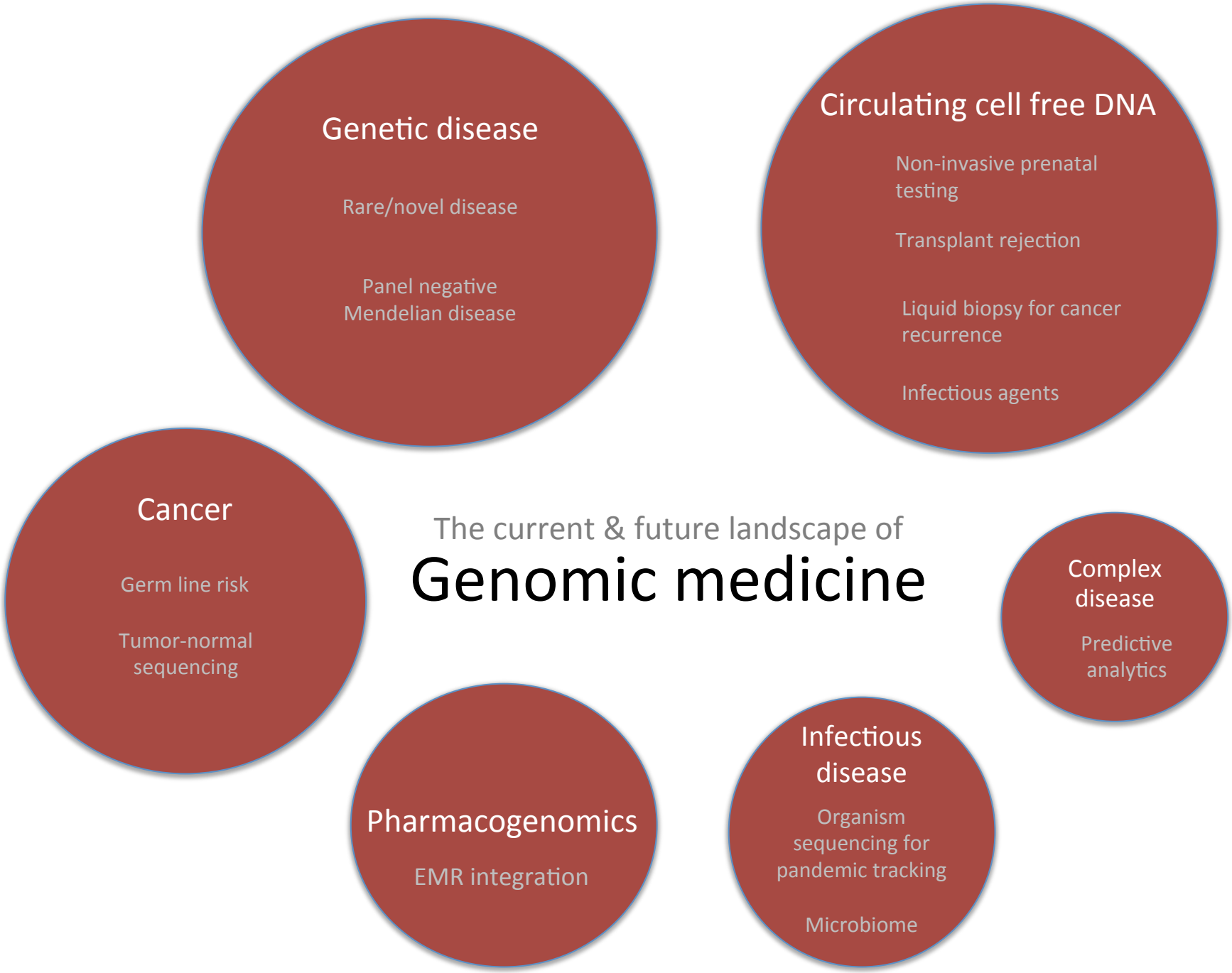


Figure 2 Size distribution of indel calls in the NA12878 trio. (**a**) Histogram of small indel calls (up to 50 bp) by size (negative size, deletion with respect to the reference sequence) for three calling algorithms. UG, UnifiedGenotyper; HC, HaplotypeCaller. (**b**) Smoothed histograms (10-bp bins) showing larger indels and peaks around ~300 bp corresponding to insertions and deletions of Alu transposable elements. Local assembly allows Platypus to detect insertions up to a few hundred basepairs in size and deletions of over 1 kb in size.

Implications

- It is possible to get to 100% coverage of many genes
- In that case, the best "test" is one that includes some genome wide coverage to allow SV detection
  - This could be long read "scaffold"
- To be cost effective, will need to have gene "augmentation" so that every base pair is callable
- This would also allow higher sensitivity for calling mosaics


Genetic disease
Rare/novel disease
Panel negative Mendelian disease
Circulating cell free DNA
Non-invasive prenatal testing
Transplant rejection
Liquid biopsy for cancer recurrence
Infectious agents
Cancer
Germ line risk
Tumor-normal sequencing
The current & future landscape of
Genomic medicine
Complex disease
Predictive analytics
Pharmacogenomics
EMR integration
Infectious disease
Organism sequencing for pandemic tracking
Microbiome

Undiagnosed Diseases Network

THE UNDIAGNOSED DISEASES NETWORK